ISSN 1561-5359. Штучний інтелект, 2019, № 3-4

УДК 004.9

О.М. Березький, О.Й. Піцун, Т.М. Долинюк, Ю.М. Батько
Тернопільський національний економічний університет, Україна
вул. Львівська, 11, м. Тернопіль, 46004

РОЗПАРАЛЕЛЕННЯ ПОШУКУ МІКРООБ’ЄКТІВ НА ЦИТОЛОГІЧНИХ ЗОБРАЖЕННЯХ ЗА ШАБЛОНОМ

O.M. Berezsky, O.Y. Pitsun, T.M. Dolynuyk, Yu.M. Batko
Ternopil National Economic University, Ukraine
11, Lvivska St., Ternopil, 46004

PARALLELING PROCESS OF SEARCHING OBJECTS ON CYTOLOGICAL IMAGES BY A TEMPLATE

Проведено аналіз сучасних підходів до пошуку елементів на зображеннях. Розроблено алгоритм пошуку мікрооб’єктів на гістологічних і цитологічних зображеннях з використанням бази даних. Розроблено ярусно-паралельну форму розпаралелення процесу пошуку мікрооб’єктів за шаблоном. Здійснено програмну реалізацію пошуку мікрооб’єктів за шаблоном. Отримані результати демонструють, що час роботи програмного модуля із застосуванням процесу розпаралелення пришвидшує процес опрацювання в середньому на 20% для цитологічних зображень.

Ключові слова: біомедичні зображення, ярусно-паралельна форма, пошук за шаблоном, цитологія, ядра клітин

Modern approaches to finding image elements are analyzed. An algorithm for searching micro-objects in histological and cytological images using a database is developed. A tiered-parallel form of parallelization of the process of micro-object pattern search is designed. Micro-object pattern search software is implemented. The obtained results show that the operating time of the software module with parallelization speeds up the processing on average by 20% for cytological images.

Keywords: biomedical images, tiered-parallel shape, template matching, cytology, cell nucleus

Вступ

На даний час у медицині для діагностування ракових та передракових станів молочної залози широко використовують гістологічний і цитологічний аналіз [1]. Із розвитком апаратних та програмних засобів, із застосування засобів штучного інтелекту та комп’ютерного зору, значно збільшилась кількість програмних засобів для аналізу зображень. Використання сучасних IT-технологій значно підвищує продуктивність та якість роботи лікарів-діагностів.

Форма та взаєморозміщення ядер клітин на цитологічних зображеннях може вказувати на наявність певних патологій, які потребують детальнішого дослідження. Однак, зображення, отримані в результаті мікроскопічного дослідження клітин, характеризуються високим рівнем різного роду шумів, що ускладнює процес автоматичної обробки та виділення ядер клітин. До того ж, ядра клітин у більшості випадків не є чітко виділеними, а навпаки – розмитими та такими, що дотикаються одне до одного.

Однією з технологій пошуку елементів на зображенні є пошук елементів за наперед заданим шаблоном. Пошук елементів за шаблоном – це техніка цифрової обробки зображень для пошуку невеликих частин зображення, які відповідають заданому зразку. У більшості випадків лікар обробляє стек цитологічних зображень, отриманих з камери мікроскопа. Кожне зображення характеризується високою роздільною здатністю, тому пошук навіть декількох наперед заданих шаблонів на вхідних зображеннях може значно завантажити робочу станцію лікаря-діагноста.

Для пришвидшення процесу пошуку елементів за шаблоном у даній статті розроблено основні кроки розпаралелення даного процесу.

Постановка проблеми

Отже, проблемою, що існує, є розроблення та впровадження засобів пошуку елементів на зображенні за наперед заданим шаблоном та пришвидшення процесу їх обчислення, формування бази шаблонів патологічних станів. Це дозволить значно підвищити якість досліджень та спростити процес постановки діагнозу.

Аналіз публікацій

Науковці значну увагу приділяють автоматизації процесу діагностування передракових та ракових станів. Оскільки одним із основних способів діагностування онкологічних захворювань є аналіз зображень зрізів тканин, то значна увага приділяється засобам розпізнавання та класифікації зображень. Наприклад, автори [2,3] пропонують алгоритм виділення ядер клітин та мікроскопічних зображень на основі їх форми та забарвлення.

У роботі [4] наведено метод пошуку об’єктів на зображенні, базуючись на афінних перетвореннях та максимальній точці кореляції для подальшого максимального накладання.

У статтях [5,6] наведено аналіз засобів для розпаралелення процесу виділення контурів ядер клітин на широкоформатних зображеннях за допомогою графічних процесорів. Автори у статті [7] наводять порівняльний аналіз процесу опрацювання біомедичних зображень із та без використання розпаралелення на графічних процесорах, що доводить актуальність застосування різних підходів до розпаралелення обчислень для пришвидшення процесу діагностування.

Важливим етапом у процесі пошуку об’єктів за шаблоном є підготовка зображення – попереднє оброблення, фільтрація, гістограмне вирівнювання. Автори Petr Strakos та Milan Jaros у роботі [8,9] наводять різні підходи та результати розпаралелення процесу фільтрації біомедичних зображень.

У статтях [10,11] наведено порівняльний аналіз існуючих програмно-апаратних засобів для автоматизації процесу діагностування гістологічних та цитологічних зображень. Наявність алгоритмів розпаралелення є значною перевагою.

Мета дослідження

Отже, метою роботи є розробка узагальненої структури модуля пошуку елементів на цитологічному зображенні за шаблоном та його розпаралелення для системи «HIAMS» [12] для підвищення швидкості та точності діагностування ракових захворювань.

Узагальнена структура модуля аналізу зображень

Модуль аналізу зображень є основною частиною програмної системи, який здійснює такі функції: завантаження стеку вхідних зображень у пам’ять, попередньої обробки, фільтрації, завантаження наперед заданих шаблонів, пошук об’єктів за ними і вивід кінцевих результатів. Структура модуля наведена на рисунку 1.

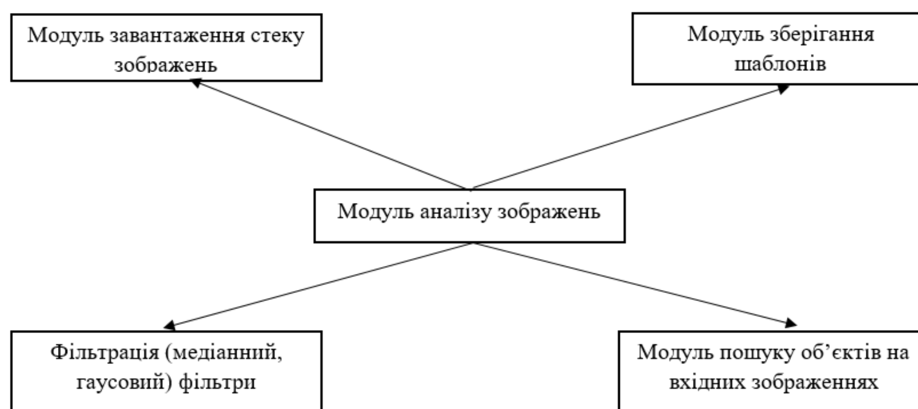


Рис. 1. Структура модуля аналізу зображень

Модуль завантаження у пам’ять отримує файли вхідних зображень, на яких буде проводитися пошук об’єктів, та завантажує їх у базу даних програми. Завантаження зображень проводиться як із внутрішнього сховища комп’ютера, так і з зовнішніх носіїв чи мережі Інтернет.

Після завантаження вхідних зображень відбувається їхня попередня обробка, а саме: фільтрація медіанним або гаусовим фільтром. Небажаний ефект від завад усувається фільтрацією.

Розпаралелення процесу пошуку об’єктів на зображенні

Пошук за шаблоном на декількох зображеннях є незалежною задачею, тому доцільно застосовувати процес розпаралелення. Ярусно-паралельну форму представлення розпаралелення процесу пошуку елементів за шаблоном наведено на рисунку 2.

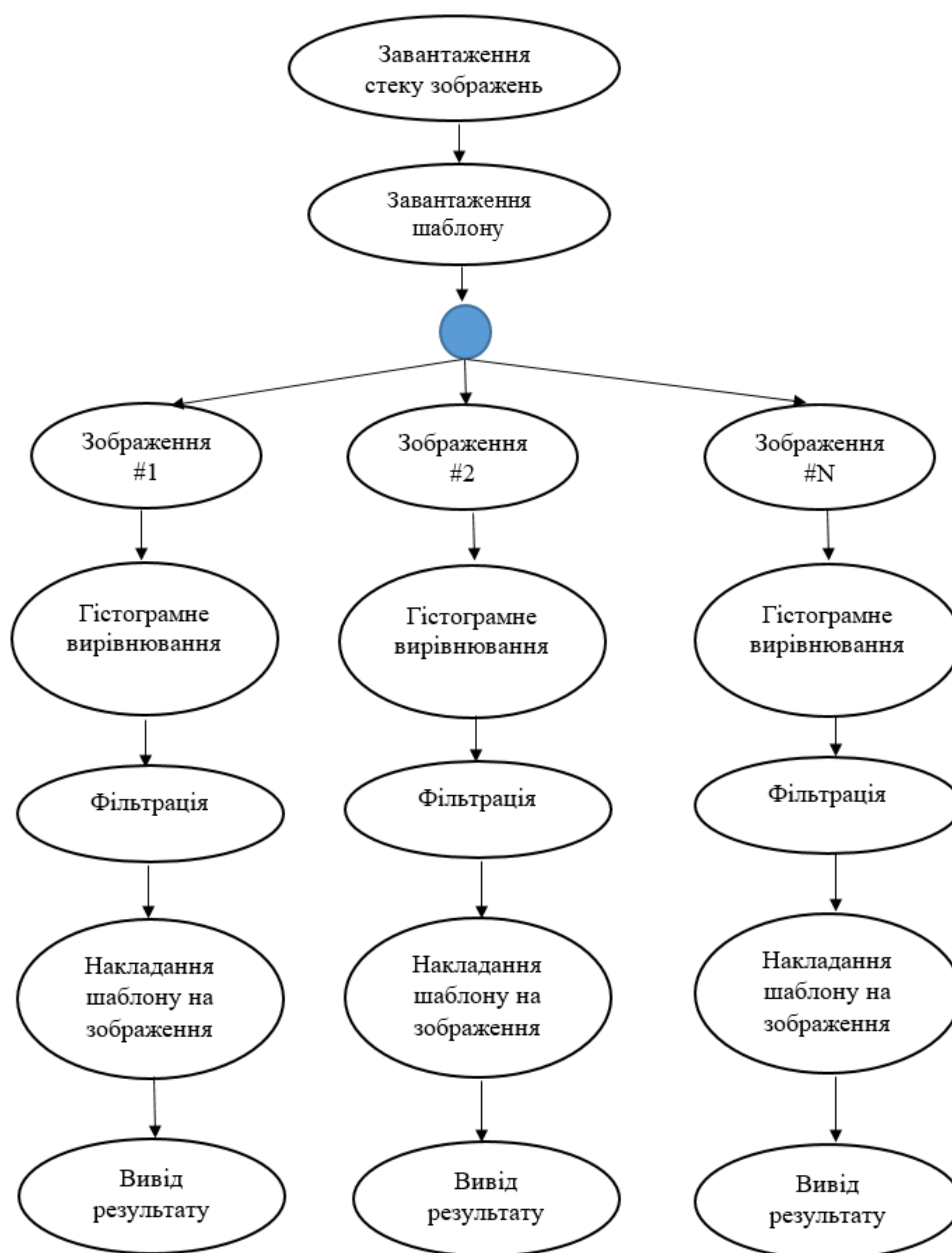


Рис. 2. Ярусно-паралельна форма представлення розпаралелення процесу пошуку елементів за шаблоном

На першому етапі відбувається завантаження шуканого шаблону в пам'ять комп'ютера. На наступному – проходить розгалуження процесу обробки зображень, оскільки дані етапи є незалежними і, відповідно, їх можна розподілити на окремі потоки. У результаті роботи на виході отримується нове зображення і, якщо на ньому присутні подібні ядра клітин, то вони будуть відповідно виділені. Надалі лікар-діагност матиме можливість детальніше звернути увагу на ті цитологічні зображення, на яких знайдено відповідні області.

Алгоритм пошуку мікрооб'єктів за шаблоном

Алгоритм пошуку мікрооб'єктів за шаблоном складається з таких кроків:

1) Виділення контурів ядер клітин за алгоритмом, розробленим у статті [13].
2) Ідентифікація трьох точок на контурах ядер клітин і обчислення афінних коефіцієнтів [14].

Коефіцієнти афінних перетворень, що описують зміну форми і розміщення початкового ядра, отримуються із системи рівнянь:

$$\begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ e \end{bmatrix} = \begin{bmatrix} \widetilde{x}_1 \\ \widetilde{x}_2 \\ \widetilde{x}_3 \end{bmatrix},$$

$$\begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{bmatrix} \cdot \begin{bmatrix} c \\ d \\ f \end{bmatrix} = \begin{bmatrix} \widetilde{y}_1 \\ \widetilde{y}_2 \\ \widetilde{y}_3 \end{bmatrix},$$

де (x_1, y_1), (x_2, y_2), (x_3, y_3) – точки, що належать першому ядру;

$(\widetilde{x}_1, \widetilde{y}_1)$, $(\widetilde{x}_2, \widetilde{y}_2)$, $(\widetilde{x}_3, \widetilde{y}_3)$ – точки, що належать другому ядру;
a, b, c, d, e, f – коефіцієнти афінного перетворення, що відображають перше ядро і друге.
Невідомі коефіцієнти визначаються так:

$$a = \frac{y_1(\widetilde{x}_3 - \widetilde{x}_2) + y_2(\widetilde{x}_1 - \widetilde{x}_3) + y_3(\widetilde{x}_2 - \widetilde{x}_1)}{y_1(x_3 - x_2) + y_2(x_1 - x_3) + y_3(x_2 - x_1)},$$

$$b = \frac{x_1(\widetilde{x}_2 - \widetilde{x}_3) + x_2(\widetilde{x}_3 - \widetilde{x}_1) + x_3(\widetilde{x}_1 - \widetilde{x}_2)}{y_1(x_3 - x_2) + y_2(x_1 - x_3) + y_3(x_2 - x_1)},$$

$$e = \frac{x_1(y_2\widetilde{x}_3 - y_3\widetilde{x}_2) + x_2(y_3\widetilde{x}_1 - y_1\widetilde{x}_3) + x_3(y_1\widetilde{x}_2 - y_2\widetilde{x}_1)}{y_1(x_3 - x_2) + y_2(x_1 - x_3) + y_3(x_2 - x_1)},$$

$$c = \frac{y_1(\widetilde{y}_3 - \widetilde{y}_2) + y_2(\widetilde{y}_1 - \widetilde{y}_3) + y_3(\widetilde{y}_2 - \widetilde{y}_1)}{y_1(x_3 - x_2) + y_2(x_1 - x_3) + y_3(x_2 - x_1)},$$

$$d = \frac{x_1(\widetilde{y}_2 - \widetilde{y}_3) + x_2(\widetilde{y}_3 - \widetilde{y}_1) + x_3(\widetilde{y}_1 - \widetilde{y}_2)}{y_1(x_3 - x_2) + y_2(x_1 - x_3) + y_3(x_2 - x_1)},$$

$$f = \frac{x_1(y_2\widetilde{y}_3 - y_3\widetilde{y}_2) + x_2(y_3\widetilde{y}_1 - y_1\widetilde{y}_3) + x_3(y_1\widetilde{y}_2 - y_2\widetilde{y}_1)}{y_1(x_3 - x_2) + y_2(x_1 - x_3) + y_3(x_2 - x_1)}.$$

3) Пересування шаблону по усіх можливих позиціях та розрахунок для кожної позиції «подібності» між шаблоном і областю на зображенні.
4) Обчислення значення найбільшої подібності здійснюється згідно з формулою:

$$cor = \frac{\sum_{i=0}^{N-1}\left(x_i - \bar{x}\right) * \left(y_i - \bar{y}\right)}{\sqrt{\sum_{i=0}^{N-1}\left(x_i - \bar{x}\right)^2 \sum_{i=0}^{N-1}\left(y_i - \bar{y}\right)^2}},$$

де x_i – значення пікселя шаблонного зображення,

$\overline{x}$ – середнє значення пікселів у шаблонному зображенні,

y_i – значення пікселя вхідного зображення,

$\overline{y}$ – середнє значення пікселів вхідного зображення,

N – загальна кількість пікселів.

Комп'ютерні експерименти

Для тестування роботи розробленого програмного модуля було оброблено цитологічні зображення із бази даних «BPCI2100» [15]. Вибірку зображень поділено на такі класи: (1) цито-фібро-кістозна мастопатія; (2) цито-кістозна мастопатія; (3) цито-непроліферативна фібромастопатія; (4) цито-непроліферативна мастопатія.

Розроблений модуль є частиною системи автоматизованої мікроскопії «HIAMS». Результат роботи програми наведено на рисунку 3.

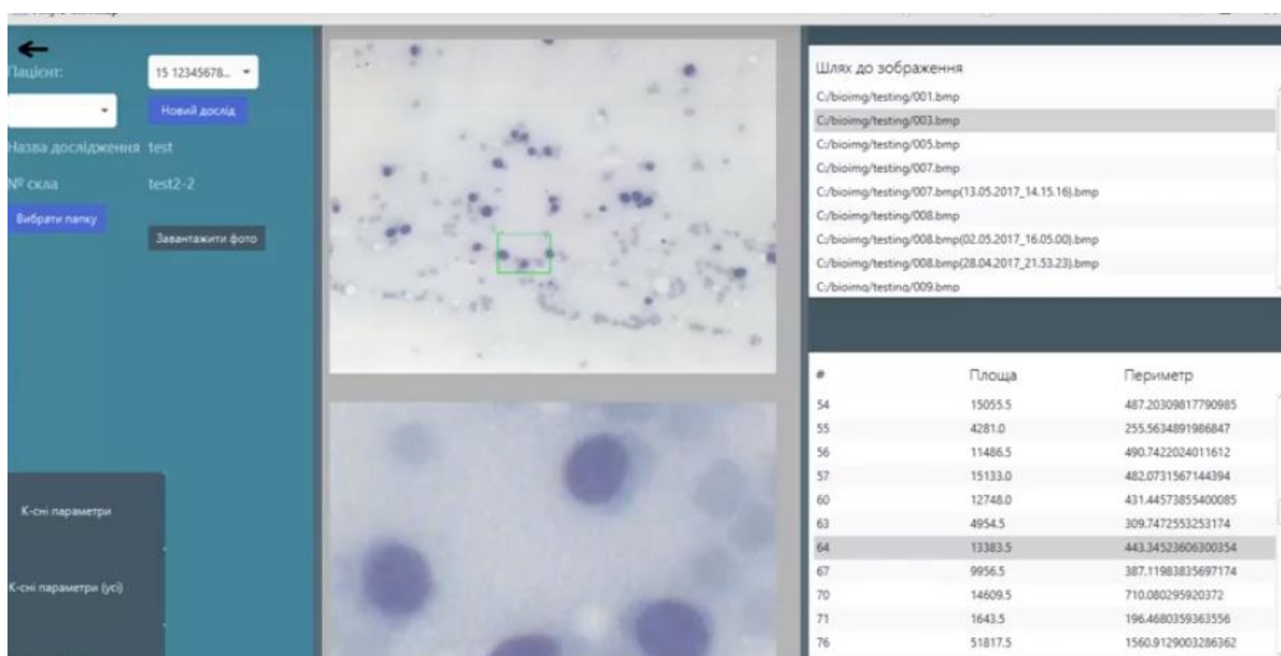
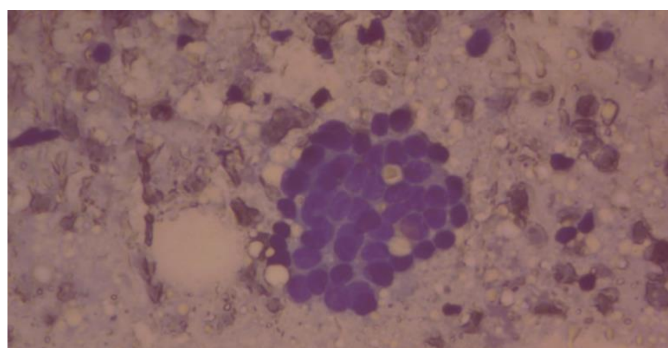


Рис. 3. Графічний інтерфейс вікна системи автоматизованої мікроскопії «HIAMS»

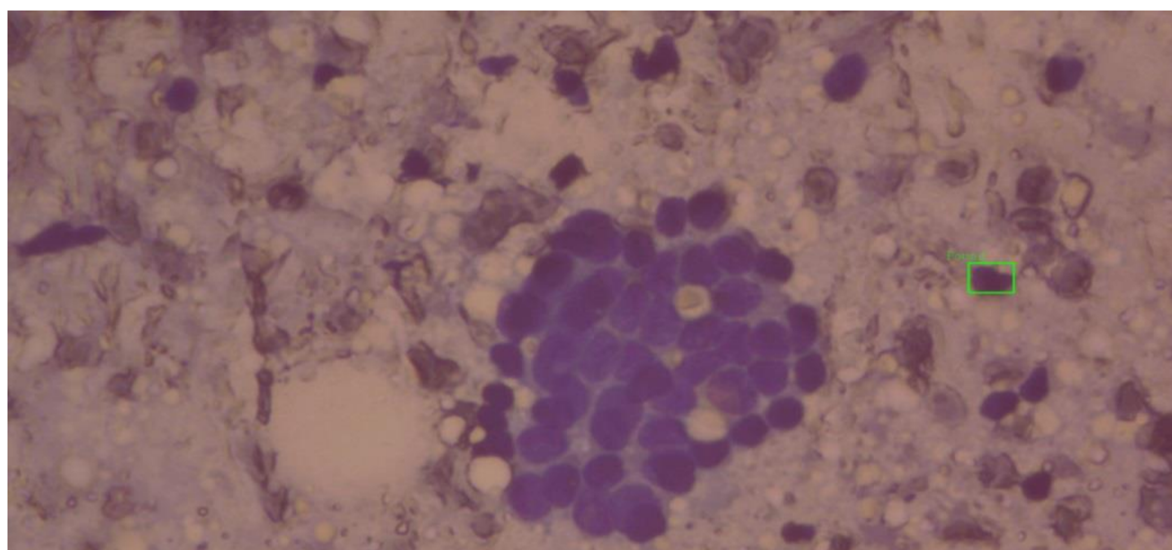
На рисунку 4 наведено приклад вхідного зображення та шаблонного зображення. Результат пошуку об'єкта наведено на рисунку 4(в).



а) вхідне зображення



б) шаблон



в) результат

Рис. 4. Пошук об’єкта на зображенні за шаблоном

Результати роботи програмного модуля для пошуку об’єкта на 20-и вхідних зображеннях розміром 3000 x 3000 пікселів наведено на рисунку 5.

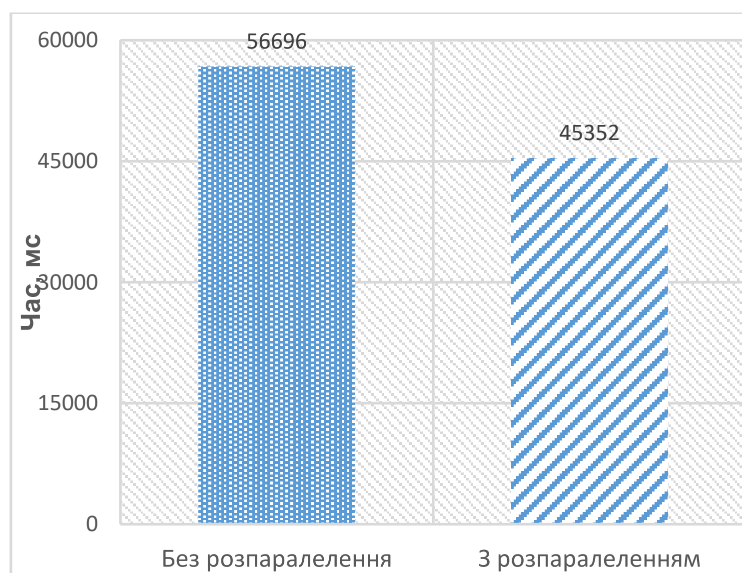


Рис. 5. Час пошуку об’єктів за шаблоном для стеку із 20-и зображень

Як видно з рисунку 5, час роботи програмного модуля з пошуку об'єктів за шаблоном на 20-ьох широкоформатних зображеннях приблизно на 20% менший, ніж без використання розпаралелення. Було проведено 5 спроб для кожного випадку і, так як час роботи програм був різний, на графіку відображено середній час для кожного випадку. Даний результат підтверджує актуальність задачі розпаралелення процесів опрацювання зображень.

Висновки

У запропонованій роботі розроблено алгоритм пошуку мікрооб'єктів за шаблоном та ярусно-паралельну форму процесу пошуку мікрооб'єктів на зображенні за шаблоном. Отримані результати показують, що ефективність застосування процесу розпаралелення складає близько 20% для цитологічних зображень із бази даних «BPCI2100».

Література

1. Методи, алгоритми і програмні засоби опрацювання біомедичних зображень / Березький О.М., Батько Ю.М., Березька К.М., Вербовий С.О., Дацко Т.В., Дубчак Л.О., Ігнатєв І.В., Мельник Г.М., Николюк В.Д., Піцун О.Й. – Тернопіль: Економічна думка, ТНЕУ, 2017. – 350 с.
2. Zarella M. Automated gland and nuclei segmentation for grading of prostate and breast cancer histopathology / M.D. Zarella, Garcia F., Breen D. // ICBBT '17 Proceedings of the 9th International Conference on Bioinformatics and Biomedical Technology. – 2017. – pp. 11-15. doi10.1145/3093293.3093307
3. Zarella, M An optimized color transformation for the analysis of digital images of hematoxylin / M. Zarella, D. Breen, A. Plagov, F. Garcia // Journal of Pathology Informatics, 6(1):33, 2015.
4. D´ıaz G. Automatic Clump Splitting for Cell Quantification in Microscopical Images / G. D´ıaz, F. Gonzalez, E. Romero // Springer-Verlag Berlin Heidelberg, CIARP 2007, LNCS 4756, pp. 763–772, 2007.
5. Lotfollahi M. Three-dimensional GPU-accelerated active contours for automated localization of cells in large images / M. Lotfollahi, S. Berisha, L. Saadatifard, L. Montier, J. Žiburkus, D. Mayerich // PLoS ONE 14(6): e0215843. https://doi.org/10.1371/journal.pone.0215843
6. Irshad H Methods for nuclei detection, segmentation, and classification in digital histopathology: a review—current status and future potential / H. Irshad, A. Veillard, L. Roux, D. Racoceanu // IEEE reviews in biomedical engineering. 2014;7:97–114. pmid:24802905
7. Berezsky O. GPU – based biomedical image processing / O. Berezsky, O. Pitsun, L. Dubchak, P. Lyaschynsky // Proceedings of XIV International Conference Perspective Technologies and methods in mems design (MEMSTECH 2018) 18-22 April, 2018, Lviv-Polyana, Ukraine, pp. 96-99.
8. Strakos P. Parallel Implementation of Collaborative Filtering Technique for Denoising of CT Images / P. Strakos, M. Jaros, T. Karasek, T. Kozubek // International Conference on High Performance Computing in Science and Engineering HPCSE 2015: High Performance Computing in Science and Engineering pp 126-140.
9. Strakos P. Parallelization of the image segmentation algorithm for Intel Xeon Phi with application in medical imaging. / P. Strakos, M. Jaros, T. Karasek, L. Riha // 4th International Conference on Parallel, Distributed, Grid and Cloud Computing for Engineering. Civil-Comp Press, Stirlingshire (2015)
10. Березький О.М. Інтелектуальна система автоматизованої мікроскопії аналізу гістологічних та цитологічних зображень / О.М. Березький, О.Й. Піцун, П.Б. Лящинський, Г.М. Мельник // Штучний інтелект, Київ, 2017. - №2 (76). - С. 128-140.
11. Berezsky O. Modern Automated Microscopy Systems in Oncology / O.M. Berezsky, O. Pitsun, N. Batryn, T. Datsko, K. Berezska, L. Dubchak // Proceedings of the 1st International Workshop on Informatics & Data-Driven Medicine, Lviv, Ukraine, 28-30 november 2018 – p. 16.
12. Свідоцтво про реєстрацію авторського права на твір №75360. Комп'ютерна програма «Інтелектуальна система діагностування передракових станів молочної залози на основі аналізу гістологічних та цитологічних зображень "HIAMS"». / О.М. Березький, О.Й. Піцун, Г.М. Мельник, П.Б. Лящинський. Дата реєстрації 14.12.2017 р.
13. Березький О. Аналіз алгоритмів виділення контурів зображень / О. Березький, Ю. Батько // Вісник Національного університету «Львівська політехніка». Комп'ютерні науки та інформаційні технології. – 2006. – № 565. – С. 212-216.
14. Березький О.М. Аналіз і синтез зображень пухлинних клітин / О.М. Березький// Інформаційні технології і системи. – 2005. – Т.8, №1. – С.76-85.
15. Свідоцтво про реєстрацію авторського права на твір №75359. База даних цифрових гістологічних та цитологічних зображень передракових та ракових станів молочної залози «BPCI2100». / О.М. Березький, Г.М. Мельник, С.О. Вербовий, О.Й. Піцун, В.Д. Николюк, Т.В. Дацко. Дата реєстрації 14.12.2017 р.

References

1. Berezky O.M. Methods, algorithms and software tools for the processing of biomedical images / Berezky O.M., Bat'ko Yu.M., Berezka K.M., Verbovy S.O., Datsko T.V., Dubchak L.O., Ignat'ev I.V., Melnyk G.M., Nikoluk V.D., Pitsun O.Y. - Ternopil: Economic Thought, TNEU, 2017. - 330 p.
2. Zarella M. Automated gland and nuclei segmentation for grading of prostate and breast cancer histopathology / M.D. Zarella, Garcia F., Breen D. // ICBBT '17 Proceedings of the 9th International Conference on Bioinformatics and Biomedical Technology. – 2017. – pp. 11-15. doi10.1145/3093293.3093307
3. Zarella, M. An optimized color transformation for the analysis of digital images of hematoxylin / M. Zarella, D.Breen, A. Plagov, F. Garcia // Journal of Pathology Informatics, 6(1):33, 2015.
4. D´ıaz G. Automatic Clump Splitting for Cell Quantification in Microscopical Images / G. D´ıaz, F. Gonzalez, E. Romero // Springer-Verlag Berlin Heidelberg, CIARP 2007, LNCS 4756, pp. 763–772, 2007.
5. Lotfollahi M. Three-dimensional GPU-accelerated active contours for automated localization of cells in large images / M. Lotfollahi, S. Berisha, L. Saadatifard, L. Montier, J. Žiburkus, D. Mayerich // PLoS ONE 14(6): e0215843. https://doi.org/10.1371/journal.pone.0215843
6. Irshad H Methods for nuclei detection, segmentation, and classification in digital histopathology: a review—current status and future potential / H. Irshad, A. Veillard, L. Roux, D. Racoceanu // IEEE reviews in biomedical engineering. 2014;7:97–114. pmid:24802905
7. Berezsky O. GPU – based biomedical image processing / O. Berezsky, O. Pitsun, L. Dubchak, P. Lyaschynsky // Proceedings of XIV International Conference Perspective Technologies and methods in mems design (MEMSTECH 2018) 18-22 April, 2018, Lviv-Polyana, Ukraine, pp. 96-99-
8. Strakos P. Parallel Implementation of Collaborative Filtering Technique for Denoising of CT Images / P. Strakos, M. Jaros, T. Karasek, T. Kozubek // International Conference on High Performance Computing in Science and Engineering HPCSE 2015: High Performance Computing in Science and Engineering pp 126-140.
9. Strakos P. Parallelization of the image segmentation algorithm for Intel Xeon Phi with application in medical imaging. / P. Strakos, M.Jaros, T. Karasek, L. Riha // 4th International Conference on Parallel, Distributed, Grid and Cloud Computing for Engineering. Civil-Comp Press, Stirlingshire (2015).
10. Berezsky O. Intelektualna systema avtomatyzovanoyi mikroskopiyi analizu gistologichnych ta cytologichnych zobrazhen` / O.M. Berezsky, O.Y. Pitsun, P.B. Lyashhynskyj, G.M. Melnyk // Shtuchnyj intelekt, Kyiv, 2017. - #2 (76). - C. 128-140.
11. Berezsky O. Modern Automated Microscopy Systems in Oncology / O.M. Berezsky, O. Pitsun, N. Batryn, T. Datsko, K. Berezska, L. Dubchak // Proceedings of the 1st International Workshop on Informatics & Data-Driven Medicine, Lviv, Ukraine, 28-30 november 2018 – p. 16.
12. Svidotstvo pro reiestratsiiu avtorskoho prava na tvir №75360. Kompiuterna prohrama «Intelektualna systema diahnostuvannia peredrakovykh staniv molochnoi zalozy na osnovi analizu histolohichnykh ta tsytolohichnykh zobrazhen "HIAMS"». / O.M. Berezkyi, O.Y. Pitsun, H.M. Melnyk, P.B. Liashchynskyi, P.B. Liashchynskyi. Data reiestratsii 14.12.2017.
13. Berezkyi O. Analiz alhorytmiv vydilennia konturiv zobrazhen / O. Berezkyi, Yu. Batko // Visnyk Natsionalnoho universytetu «Lvivska politekhnika». Kompiuterni nauky ta informatsiini tekhnolohii. – 2006. – № 565. – S. 212-216.
14. Berezkyi O.M. Analiz i syntez zobrazhen pukhlynnykh klityn / O.M. Berezkyi// Informatsiini tekhnolohii i systemy. – 2005. – T.8, №1. – S.76-85.
15. Svidotstvo pro reiestratsiiu avtorskoho prava na tvir №75359. Baza danykh tsyfrovykh histolohichnykh ta tsytolohichnykh zobrazhen peredrakovykh ta rakovykh staniv molochnoi zalozy «BPCI2100». / O.M. Berezkyi, H.M. Melnyk, S.O. Verbovyi, O.I. Pitsun, V.D. Nykoliuk, T.V. Datsko. Data reiestratsii 14.12.2017.

RESUME

O.M. Berezsky, O.Y. Pitsun, T.M. Dolynuyk, Yu.M. Batko

Paralleling of the process of searching objects on cytological images by the template

The use of modern IT technologies in medicine, which significantly increase the efficiency and quality of the diagnosticians' work, is analyzed. Diagnosing pathological processes in oncology is based on the analysis of biomedical (cytological and histological) images. Generally, cytological and histological images are characterized by high complexity processing. The main disadvantages are the presence of different kinds of noise and high definition. Cytological images are analyzed. It is shown that the

shape and interposition of cell nuclei in cytological images may indicate the presence of certain pathologies. The technology of image elements pattern search is used.

Image recognition and classification tools are analyzed. Classification of images relates the input biomedical image to a particular class of pathologies. In addition, the analysis of means of parallelization of the process of cell nuclei contour selection in widescreen images using graphic processors was carried out. It is shown that the parallelization of calculations accelerates the process of diagnosis. An image analysis module is developed that performs the following functions: uploading a stack of input images into memory, pre-processing, filtering, uploading predefined patterns, pattern searching for objects and outputting the final results.

An algorithm for micro-object pattern search is developed which consists of the following steps: cell nuclei contour selection, identification of three points in cell nuclei contours, calculation of affinity coefficients and calculation of the highest similarity of images. The database of histological and cytological images of precancerous and cancerous conditions of the mammary gland «BPCI2100» was used for the experiments and it was also developed by the authors. A software module was developed and tested using cytological images from the BPCI2100 database. The developed module is a component of HIAMS automated microscopy system. The obtained results show that the efficiency of the parallelization process is about 20% for cytological images from the BPCI2100 database.

Надійшла до редакції 30.05.2019